

①⑨ RÉPUBLIQUE FRANÇAISE

INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

①⑪ N° de publication :
(à n'utiliser que pour les
commandes de reproduction)

2 812 736

②① N° d'enregistrement national : 00 10353

⑤① Int Cl⁷ : G 06 F 12/16, G 06 F 13/14

①⑫

DEMANDE DE BREVET D'INVENTION

A1

②② Date de dépôt : 04.08.00.

③⑦ Priorité :

④③ Date de mise à la disposition du public de la
demande : 08.02.02 Bulletin 02/06.

⑤⑥ Liste des documents cités dans le rapport de
recherche préliminaire : *Se reporter à la fin du
présent fascicule*

⑥⑦ Références à d'autres documents nationaux
apparentés :

⑦① Demandeur(s) : NEAR TECHNOLOGIES Société ano-
nyme — FR.

⑦② Inventeur(s) : AZAMBRE HUBERT, DELBOSC JEAN
MARC et BLANCHET CLAUDE.

⑦③ Titulaire(s) :

⑦④ Mandataire(s) : BREESE MAJEROWICZ SIMONNOT.

⑤④ SYSTEME DE STOCKAGE VIRTUEL.

⑤⑦ La présente invention concerne un système pour la
sauvegarde de données provenant d'un ordinateur hôte
[mainframe] caractérisé en ce qu'il comprend un équipe-
ment informatique comprenant une interface d'entrée sortie
pour l'échange de données avec l'ordinateur hôte, ladite in-
terface comprenant un émulateur de lecteur-enregistreur de
sauvegarde, au moins un dispositif de stockage intermédiaire
et un lecteur-enregistreur à bande, l'équipement compre-
nant en outre un processeur pour transférer entre l'interface
d'entrée-sortie ou du dispositif de stockage intermédiaire et
le lecteur-enregistreur à bande, le système comprenant en
outre un superviseur comprenant une mémoire pour l'enre-
gistrement des informations relatives aux enregistrements
sur bande de l'équipement informatique, et pour comman-
der ledit équipement informatique en fonction des instruc-
tions provenant de l'ordinateur hôte.

FR 2 812 736 - A1



SYSTEME DE STOCKAGE VIRTUEL

La présente invention concerne le domaine du stockage de données informatiques sur des supports tels que des cassettes de grande capacité, par un équipement distant comprenant généralement un robot de manipulation de cassettes.

A titre d'exemple le brevet PCT WO9844423 décrit un réseau informatique comprenant un certain nombre d'unités de commande de stockage, chacune étant couplée à une pluralité d'ensembles de stockage, lesquels ensembles comprennent au moins un dispositif à mémoire de grande capacité (MSD). Chaque unité de commande de stockage peut être couplée à au moins un système de traitement hôte et à au moins une autre unité de commande de stockage de façon à commander l'accès des systèmes de traitement hôte aux dispositifs à mémoire de grande capacité. Plusieurs copies de données sont stockées dans des ensembles de stockage géographiquement distants les uns des autres, de sorte que n'importe quel hôte peut accéder à n'importe quelle copie. Chaque unité de commande de stockage comprend une interface avec un hôte qui émule un dispositif à mémoire de grande capacité indépendant du type de dispositif de stockage et une interface avec un ensemble local de stockage qui émule un hôte indépendant du type de l'hôte. Des hôtes accèdent aux données stockées au moyen d'un dressage virtuel. Les unités de commande de stockage assurent des sauvegardes et des corrections d'erreur automatiques ainsi qu'une protection en écriture des copies de sauvegarde.

Le but de l'invention est de proposer un système de sauvegarde amélioré permettant l'utilisation par un parc hétérogène d'ordinateurs hôtes, relié à un équipement de sauvegarde commun non spécifique. A cet

effet, l'invention concerne selon son acception la plus générale un système pour la sauvegarde de données provenant d'un ordinateur hôte [mainframe] caractérisé en ce qu'il comprend un équipement informatique comprenant une interface d'entrée sortie pour l'échange de données avec l'ordinateur hôte, ladite interface comprenant un émulateur de lecteur-enregistreur de sauvegarde, au moins un disque dur et un lecteur-enregistreur à bande, l'équipement comprenant en outre un processeur pour transférer entre l'interface d'entrée-sortie ou du lecteur de bande et le lecteur-enregistreur à bande, le système comprenant en outre un superviseur comprenant une mémoire pour l'enregistrement des informations relatives aux enregistrements sur bande de l'équipement informatique, et pour commander ledit équipement informatique en fonction des instructions provenant de l'ordinateur hôte.

Avantageusement, l'émulateur est constitué par un calculateur pour procéder à l'analyse des signaux provenant de l'ordinateur hôte et pour la génération d'une réponse correspondant au type de lecteur-enregistreur de cassette simulé.

L'invention concerne également un procédé pour la sauvegarde de données provenant d'un ordinateur hôte [mainframe] caractérisé en ce que l'on émule l'interface d'entrée sortie d'un équipement de sauvegarde afin qu'un équipement de sauvegarde présente, pour une machine hôte, un comportement identique à un dérouleur de bandes, ledit équipement de sauvegarde comportant un moyen de stockage intermédiaire différent d'un dérouleur de bande.

La présente invention sera mieux comprise à la lecture de la description d'un exemple non limitatif de réalisation qui suit, ce référant aux dessins annexés où :

- la figure 1 représente le schéma de principe de l'invention.

Le système décrit dans ce qui suit permet la sauvegarde de données provenant d'un parc hétérogène de machines hôtes (1) de type «Mainframe» connectées à un réseau informatique (2) de type SCSI.

L'équipement de sauvegarde (3) comporte un ou plusieurs dérouleurs de bande (4) ou « streamer » pour la sauvegarde des données sur support magnétique.

Il est connecté au réseau par le biais d'un circuit interface (5) d'entrée-sortie émulé de façon à ce que l'équipement de sauvegarde (3) soit vu par la machine hôte sous la forme d'un dérouleur de sauvegarde du type émulé, et ce pour l'ensemble des fonctions remplies par l'équipement de sauvegarde (3). L'interface émulée assure une émulation des principaux dérouleurs de bandes connus afin de permettre un dialogue transparent entre la machine hôte et l'équipement de sauvegarde (3).

L'équipement de sauvegarde (3) comporte par ailleurs au moins un dispositif de stockage intermédiaire (9) composé dans l'exemple décrit par des disques durs RAID.

L'équipement de sauvegarde comporte pour chacun des supports de sauvegarde des initiateurs (6, 7). Un calculateur commande les différentes ressources pour assurer le transfert des données depuis l'interface d'entrée sortie (5) vers le dispositif de stockage intermédiaire (9) et inversement, ainsi que le transfert des données depuis le dispositif de stockage intermédiaire (9) vers les dérouleurs de bande (4) et inversement.

Vu depuis la machine hôte, l'équipement de sauvegarde selon l'invention répond aux spécifications suivantes :

- Il présente un comportement identique au dérouleur de bandes auquel il se substitue ;
- Il améliore la vitesse de stockage des données par l'intermédiaire d'un cache disque. Afin d'accélérer la sauvegarde et la restauration des données, celles-ci sont stockées sur une partition du disque. L'accès aux données est amélioré grâce à un métamodèle des données sauvegardées qui mémorise la cartographie des données. Ce métamodèle permet un accès direct aux données stockées séquentiellement.
- Il recopie les données sur un dérouleur. Les données sauvegardées sur la partition disque sont recopiées sur bande, en reproduisant le mécanisme d'écriture initial, par l'exploitation du modèle.
- Il assure la persistance et la cohérence des données. En fin de sauvegarde, l'équipement de sauvegarde est le garant de la persistance et de la cohérence des données sur bande et partition. De plus, il permet de décorréler le type de dérouleur amont (celui que l'on émule), de l'aval (celui que l'on pilote réellement). En amont, l'équipement de sauvegarde gère un dérouleur d'un modèle et sauvegarde les données sur un dérouleur d'un autre modèle.

L'équipement de sauvegarde (3) assure les types de connexions suivantes :

- en amont : SCSI, FC, ESCON, Bus&Tag
- en aval : SCSI, FC.

L'équipement de sauvegarde gère plusieurs connexions en amont et en aval simultanément. En conséquence, il exécute plusieurs transferts en parallèle. Chaque transfert est géré par une unité de transfert.

Une unité de transfert gère trois types de liens :

- liaison avec un système hôte
- liaison avec une partition d'un disque physique
- liaison avec le dérouleur de bande.

5 Le système comporte en outre un poste superviseur (12) connecté par des liaisons série (13, 14) d'une part à la machine hôte et d'autre part à l'équipement de sauvegarde.

10 L'émulation consiste à simuler le fonctionnement SCSI d'un dérouleur de bande vis-à-vis d'une machine hôte et à gérer les réponses SCSI aux différentes demandes de l'hôte et les transferts d'enregistrement.

15 Le poste de supervision (12) commande une base de données dans laquelle sont enregistrées les étiquettes d'identifications des données sauvegardées.

20 Les volumes de données écrits par les machines hôtes sont dans un premier temps créés dans un espace disque tampon (9). Ces volumes ont une taille maximale fixée au moment de la configuration du système de sauvegarde, généralement assez faible, de l'ordre de 250 Mo. Les volumes sont dans un deuxième temps transférés vers des cartouches, en un ou plusieurs exemplaires. Seules les données réellement significatives sont transférées sur bande. Ainsi, un volume de taille maximale 250 Mo peut par

25 exemple ne contenir que 10 Mo de données. Dans ce cas, seuls ces 10 Mo sont transférés sur bande, de façon à optimiser l'espace bande.

30 L'équipement de sauvegarde gère en interne, grâce à une base de données, la liste des volumes connus, en stockant un certain nombre d'informations, telles que :

- le nom du volume
- le support sur lequel il est stocké (disque, cartouche)

- la position sur le support (numéro de partition disque, ou adresses logiques de début et de fin sur cartouche)
- etc.

5 Ces informations sont essentielles pour pouvoir retrouver un volume.

Au moment du transfert des données du cache disque vers des cartouches, on ajoute en fin de transfert de chaque volume des informations privées, appelées
10 « Informations de base ». Ces informations sont uniquement écrites sur cartouches, et sont ignorées dans les transferts en sens inverse, dans le cas où un volume est transféré depuis une cartouche vers le cache disque, par exemple en vue d'une restauration par la machine hôte.
15 Elles sont donc entièrement gérées en interne et transparentes pour les machines hôtes.

Les informations de base pour un volume donné sont écrites sous la forme d'une chaîne de caractères ASCII ayant la structure suivante :

20 *⌘titre CR LF positionDébutVolume positionFinVolume tailleVolume voieLecteur *
*voieDisque partitionDisque numéroProcesseur \ codeBarre nomCartouche *
*typeCartouche tailleUtilisée tailleCartouche *
*cptChargement nomVolume étatVolume codeHost typeCodage *
*dateEcriture heureEcriture dateLecture heureLecture *
25 *dateVidage heureVidage CR LF⌘*

titre : titre indiquant sous forme abrégée la signification des principaux champs suivants.

CR : caractère ASCII de code 0x13 (hexadécimal).

30 LF : caractère ASCII de code 0x10 (hexadécimal).

positionDébutVolume : adresse logique sur la cartouche du début du volume.

positionFinVolume : adresse logique sur la
35 cartouche de la fin du volume.

tailleVolume : taille approximative du volume en Ko.

5 *voieLecteur* : numéro du lecteur (défini dans la configuration HBS) utilisé pour réaliser le transfert du volume du cache disque vers la cartouche.

voieDisque : numéro du disque (défini dans la configuration HBS) dans lequel se trouvait le volume au moment de son transfert vers la cartouche.

10 *partitionDisque* : numéro de la partition disque dans laquelle se trouvait le volume avant son transfert sur la cartouche.

numéroProcesseur : numéro du processeur utilisé pour transférer le volume du cache disque vers la cartouche.

15 *codeBarre* : code barre de la cartouche contenant le volume.

nomCartouche : nom de la cartouche, tel qu'il est déclaré sous HBS. Ce nom est indépendant du code barre.

20 *typeCartouche* : code hexadécimal indiquant le type de la cartouche. Les valeurs possibles sont les suivantes :

0x00000001L cartouche d'exploitation

0x00000010L cartouche avec accès en lecture autorisé

0x00000020L cartouche avec accès en écriture autorisé

25 0x00000080L cartouche en cours de réorganisation

0x00000100L cartouche à réorganiser

0x00000200L cartouche à ne pas réutiliser

0x00000400L cartouche bloquée vide

0x00000800L cartouche réorganisée

30 0x00001000L cartouche de type archive

0x00002000L cartouche de type miroir

0x00010000L cartouche pour lecteur DLT

0x00020000L cartouche pour lecteur Exabyte

0x00040000L cartouche pour lecteur 3480
0x00080000L cartouche pour lecteur 3590
0x01F00000L masque pour numéro du pool archive ou miroir
 auquel appartient la cartouche

5 Le code utilisé pour le champ *typeCartouche*
peut éventuellement être une combinaison des valeurs
précédentes.

tailleUtilisée : taille totale des données
stockées sur la cartouche, en Méga Octets.

10 *tailleCartouche* : capacité maximale de la
cartouche, en Méga Octets.

cptChargement : compteur de chargement de la
cartouche. Indique le nombre de fois que la cartouche a été
chargée dans un lecteur. Cette information sert à
15 déterminer l'usure de la cartouche.

nomVolume : nom du volume, tel qu'il est connu
de la machine hôte.

étatVolume : code hexadécimal indiquant l'état
du volume. Ce code est une combinaison des indicateurs dont
20 les masques d'accès et les valeurs possibles sont les
suivants :

0x00000001L 1 si le volume est valide, 0 si invalide
 (ancienne version ou volume logiquement
 effacé)

25 0x00000008L 1 si le volume est de type miroir

0x00000010L 1 si le volume a une copie miroir sur une
 autre cartouche

0x00000020L 1 si une copie de ce volume doit être faite
 sur une cartouche miroir

30 0x00001000L 1 si le volume est de type archive

0x00002000L 1 si le volume est partagé entre plusieurs
 systèmes hôtes

0x00010000L 1 si le volume doit toujours être copiés
sur des cartouches DLT

0x00020000L 1 si le volume doit toujours être copiés
sur des cartouches Exabyte

5 0x00040000L 1 si le volume doit toujours être copiés
sur des cartouches 3480

0x00080000L 1 si le volume doit toujours être copiés
sur des cartouches 3590

10 0x01F00000L numéro du pool archive ou miroir (de 0 à
31)

codeHost : numéro de la machine hôte de la
configuration HBS à laquelle appartient le volume.

typeCodage : code de caractères utilisé dans
l'entête du volume (0 = ASCII, 1 = EBCDIC).

15 *dateEcriture* : date de dernière écriture ou
modification du volume par la machine hôte, sous la forme
aaaa-mm-jj

20 *heureEcriture* : heure de dernière écriture ou
modification du volume par la machine hôte, sous la forme
hh:mm:ss

dateLecture : date de dernier accès en lecture
du volume par la machine hôte, sous la forme aaaa-mm-jj

25 *heureLecture* : heure de dernier accès en
lecture du volume par la machine hôte, sous la forme
hh:mm:ss

dateVidage : date du transfert du volume du
cache disque vers la cartouche, sous la forme jj-mm-aaaa

30 *heureVidage* : heure du transfert du volume du
cache disque vers la cartouche, sous la forme hh:mm:ss Les
informations de base sont cumulatives, pour accélérer
l'analyse des cartouches en vue d'une reconstruction de la
base de données.

V1	B1	V2	B2	V3	B3	V4	B4	V5		B5
----	----	----	----	----	----	----	----	----	--	----

Supposons qu'une bande contient les volumes V1, V2, V3, V4 et V5. Les informations de base associées à chacun de ces volumes sont appelées B1, B2, B3, B4 et B5. On a donc sur la bande : Les informations de base B1 ne contiennent que des informations concernant le volume V1. Les informations de base B2 contiennent dans un seul enregistrement de données un cumul de B1 et des informations concernant le volume V2. B2 contient donc les informations de V1 et V2.

Les informations de base B3 contiennent dans un seul enregistrement de données un cumul de B2 et des informations concernant le volume V3. B3 contient donc les informations de V1, V2 et V3.

Les dernières informations de base de la cartouche, B5 dans l'exemple précédent, contiennent donc un cumul de toutes les informations concernant tous les volumes présents sur la cartouche.

Dans le cas où une cartouche contient de très nombreux volumes, les informations de base cumulées peuvent atteindre une taille importante. Pour limiter cet accroissement, une taille maximale a été arbitrairement fixée à 132 Ko. Lorsque la construction standard des informations de base d'un volume dépasse 132 Ko, l'équipement (3) affecte à ce volume des informations de base réduites, contenant uniquement les informations de base de ce nouveau volume, sans cumul des informations des volumes précédents. Pour les volumes suivants, le mécanisme standard de cumul des informations du volume courant avec celles du volume précédent est repris.

En cas de perte totale de la base de données du système, la reconstruction complète de la base peut être

effectuée grâce à ces informations de base. Une fonctionnalité intégrée au code des processeurs permet d'analyser une cartouche pour en extraire les dernières informations de base. Cette analyse peut également être effectuée par un logiciel externe : il suffit de se positionner en fin de bande, de reculer d'un enregistrement et de lire le dernier enregistrement de données. Les informations de base ainsi récupérées en fin de cartouche contiennent une description des volumes de la cartouche. Comme décrit au paragraphe précédent, si le champ `adresseVolume` du premier volume contient une valeur différente de 0, ceci signifie que le premier volume n'est pas au début de la bande. On doit en déduire qu'il s'agit d'informations de base réduites. Il suffit alors de se positionner sur la cartouche à l'adresse `adresseVolume`, puis de reculer de 1 enregistrement pour pouvoir lire les informations de base du volume précédent. Celles-ci sont elles-mêmes un cumul des informations de base des volumes précédents.

On doit poursuivre l'analyse arrière de la cartouche jusqu'à trouver des informations de base avec un champ `adresseVolume` à 0 pour le premier volume. L'ensemble des volumes de la cartouche peut alors être connu en cumulant toutes les informations de base récupérées.

La reconstruction de la base nécessite la récupération de toutes les informations de base stockées sur toutes les cartouches de la librairie, puis leur analyse, par un logiciel approprié. L'ensemble de ces informations comprend toutes celles nécessaires pour reconstruire la base. Pour ce faire, il convient non seulement de dresser la liste de tous les volumes contenus sur toutes les cartouches, mais aussi de déterminer pour chaque volume d'une cartouche s'il est valide pour la

machine hôte ou non. En effet, un même volume (même nom, même système hôte) peut être présent sur plusieurs cartouches différentes, ou à plusieurs endroits sur une même cartouche. Ceci peut se produire pour les

5 raisons suivantes :

soit il s'agit de plusieurs versions différentes d'un même volume qui a été mis à jour plusieurs fois par la machine hôte

10 soit il s'agit des mêmes données qui ont été déplacées par HBS en interne. Dans tous les cas, l'analyse des informations de base *dateEcriture* et *heureEcriture* de toutes les occurrences de ce volume permet de déterminer laquelle est la plus récente, et donc la seule valide. Si la version la plus récente est présente à plusieurs

15 endroits (mêmes informations *dateEcriture* et *heureEcriture*), n'importe laquelle de ces occurrences peut être utilisée pour devenir la version valide du volume dans la nouvelle base. Il suffit ensuite de recréer une base de données vide, et de remplir toutes les tables avec les

20 informations collectées.

REVENDEICATIONS

1 - Système pour la sauvegarde de données
provenant d'un ordinateur hôte [mainframe] caractérisé en
ce qu'il comprend un équipement informatique comprenant une
5 interface d'entrée sortie pour l'échange de données avec
l'ordinateur hôte, ladite interface comprenant un émulateur
de lecteur-enregistreur de sauvegarde, au moins un
dispositif de stockage intermédiaire et un lecteur-
enregistreur à bande, l'équipement comprenant en outre un
10 processeur pour transférer entre l'interface d'entrée-
sortie ou du dispositif de stockage intermédiaire et le
lecteur-enregistreur à bande, le système comprenant en
outre un superviseur comprenant une mémoire pour
l'enregistrement des informations relatives aux
15 enregistrements sur bande de l'équipement informatique, et
pour commander ledit équipement informatique en fonction
des instructions provenant de l'ordinateur hôte.

2 - Système pour la sauvegarde de données selon
20 la revendication 1 caractérisé en ce que l'émulateur est
constitué par un calculateur pour procéder à l'analyse des
signaux provenant de l'ordinateur hôte et pour la
génération d'une réponse correspondant au type de lecteur-
enregistreur de cassette simulé.

25 3 - Système pour la sauvegarde de données selon
la revendication 1 ou 2 caractérisé en ce que le dispositif
de stockage intermédiaire est composé par au moins un
disque dur.

30 4 - Système pour la sauvegarde de données selon
l'une quelconque des revendications précédentes,
caractérisé en ce qu'il comporte un poste superviseur (12)

comportant une mémoire pour l'exploitation d'une base de données contenant les étiquettes d'identification des données sauvegardées.

5 5 - Système pour la sauvegarde de données selon la revendication 4 caractérisé en ce que les données numériques formant les étiquettes d'identification comprennent le nom du volume, le support sur lequel il est stocké (disque, cartouche), et la position sur le support
10 (numéro de partition disque, ou adresses logiques de début et de fin sur cartouche).

 6 - Système pour la sauvegarde de données selon l'une quelconque des revendications précédentes,
15 caractérisé en ce que le poste superviseur (12) est relié à l'équipement de sauvegarde et à la machine hôte par des liaisons série.

 7 - Système pour la sauvegarde de données selon l'une quelconque des revendications précédentes,
20 caractérisé en ce que l'équipement de sauvegarde est relié à la machine hôte par une liaison de type SCSI ou FC.

 8 - Procédé pour la sauvegarde de données provenant d'un ordinateur hôte [mainframe] caractérisé en ce que l'on émule l'interface d'entrée sortie d'un équipement de sauvegarde afin qu'un équipement de sauvegarde présente, pour une machine hôte, un comportement identique à un dérouleur de bandes, ledit équipement de sauvegarde comportant un moyen de stockage intermédiaire
30 différent d'un dérouleur de bande.

1/1

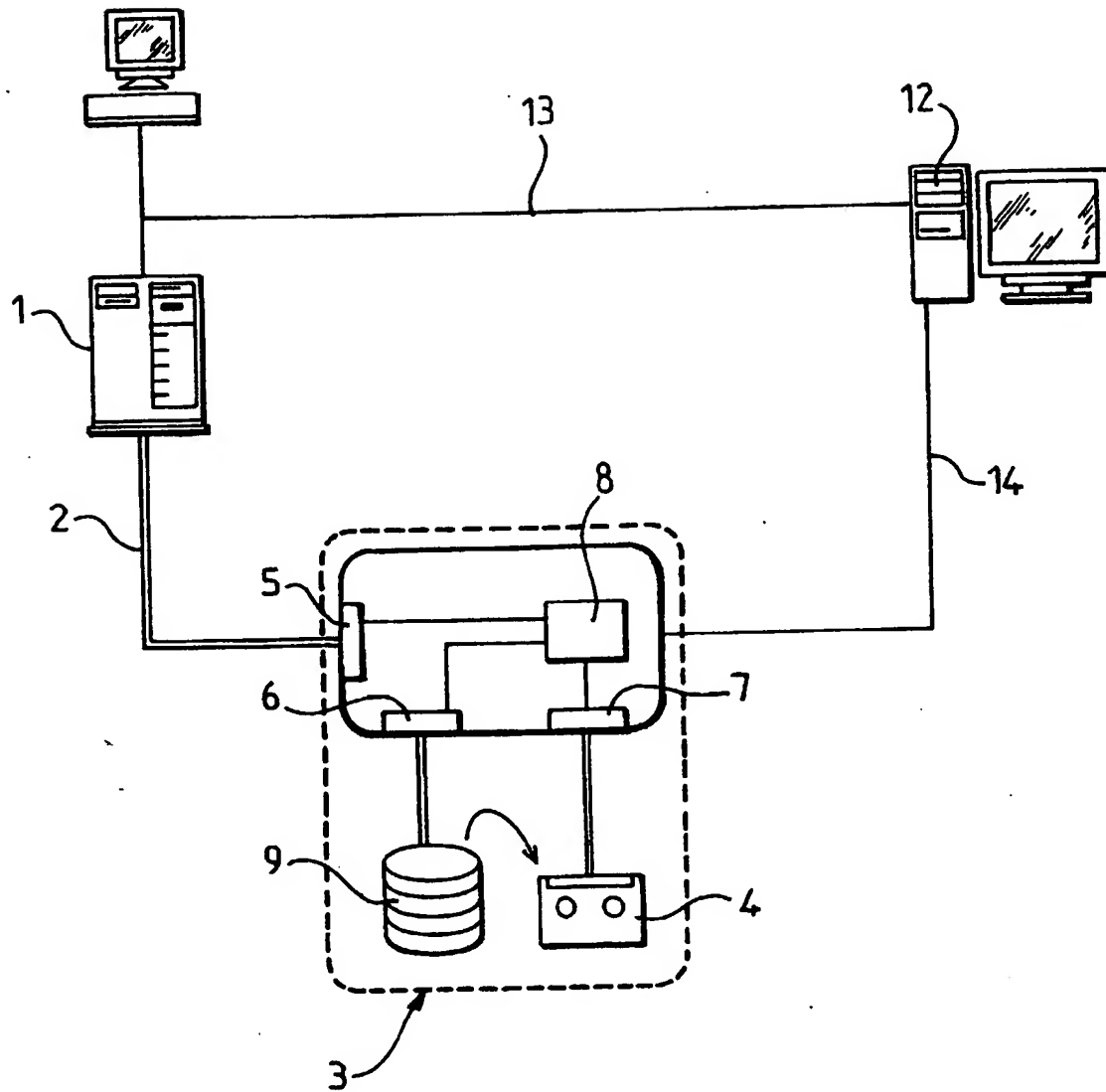


FIG.1



RAPPORT DE RECHERCHE PRÉLIMINAIRE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

2812736

N° d'enregistrement
nationalFA 591565
FR 0010353

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	US 5 809 511 A (PEAKE JONATHAN WAYNE) 15 septembre 1998 (1998-09-15)	1-3,8	G06F12/16 G06F13/14
A	* colonne 3, ligne 60 - colonne 6, ligne 24; figures 1,4 *	4-7	
D,A	WO 98 44423 A (ARK RESEARCH CORP ;BERGSTEN JAMES R (US)) 8 octobre 1998 (1998-10-08) * abrégé *	1-8	
			DOMAINES TECHNIQUES RECHERCHÉS (Int.CL.7)
			G06F
Date d'achèvement de la recherche		Examineur	
14 juin 2001		Moens, R	
CATÉGORIE DES DOCUMENTS CITÉS X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant			

THIS PAGE BLANK (USPTO)